Linear algebra

22^{nd} November 2005

Definition 1. Let A be a square, nonsingular matrix. Then the *inverse matrix* A^{-1} of A is a unique matrix for which,

$$AA^{-1} = I = A^{-1}A$$

§

Example 1. An inverse matrix may be found using the formula,

$$A^{-1} = \frac{1}{|\,A\,|}\,\mathrm{Adj}\,A$$

Example 2. Matrix equations of the form $A\mathbf{x} = \mathbf{b}$ can be solved with the help of the inverse matrix A^{-1} as $\mathbf{x} = A^{-1}\mathbf{b}$, where A is an $n \times n$ matrix, $\mathbf{x}$ a vector of size n whose components are variables, and $\mathbf{b}$ a vector of size n containing constants.

Theorem 1. Let A be the coefficient matrix and A_i a matrix formed from A by replacing the column of coefficients of x_i with the column vector of constants. Cramer's rule solves a system of linear equations through the use of determinants as follows.

$$x_i = \frac{|\,A_i\,|}{|\,A\,|}$$

§

Problem 1. Prove Cramer's rule, Theorem 1.

§

Definition 2. Let a system of n functions not necessarily linear be

$$\begin{aligned} y_1 &= f_1\,(x_1, \ldots, x_n) \\ &\vdots \\ y_n &= f_n\,(x_1, \ldots, x_n) \end{aligned}$$

Then a *Jacobian determinant* comprises all the first-order partial derivatives of the system arranged in ordered sequence, that is

$$|\,J\,| = \left|\,\frac{\partial y_1, \ldots, \partial y_n}{\partial x_1, \ldots, \partial x_n}\,\right| = \begin{vmatrix} \frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial y_n}{\partial x_1} & \cdots & \frac{\partial y_n}{\partial x_n} \end{vmatrix}$$

§

Theorem 2. Let a system of n equations be $y_i = f_i\,(x_1, \ldots, x_n)$, $i = 1, \ldots, n$. If $|\,J\,| = 0$, then y_i are functionally dependent. On the other hand if $|\,J\,| \neq 0$, then y_i are functionally independent.

§

Problem 2. Prove Theorem 2.

§

Definition 3. A determinant $|\,H\,|$ composed of all the second-order partial derivatives, with the direct partials on the principal diagonal and the cross partials off the same, is called a *Hessian*. In other words, let a multivariable function be $z = f(x, y)$. Then the Hessian of z is

$$|\,H\,| = \begin{vmatrix} z_{xx} & z_{xy} \\ z_{yx} & z_{yy} \end{vmatrix}$$

where $z_{xy} = z_{yx}$. Moreover, the *first principal minor* is $|\,H_1\,| = z_{xx}$ and the *second principal minor* is

$$|\,H_2\,| = \begin{vmatrix} z_{xx} & z_{xy} \\ z_{xy} & z_{yy} \end{vmatrix} = z_{xx}z_{yy} - (z_{xy})^2$$

§

Theorem 3. Let a multivariable function be $z = f(x, y)$, and let the first-order conditions $z_x = z_y = 0$ are met. Then a sufficient condition for z to be at optimum is $z_{xx}z_{yy} > (z_{xy})^2$ together with $z_{xx}, z_{yy} < 0$ in case of a maximum and $z_{xx}, z_{yy} > 0$ in case of a minimum.

§

Problem 3. Prove Theorem 3, the theorem on the optimality of a multivariable function.

§

Definition 4. From Definition 3, if $|\, H_1 \,| > 0$ and $|\, H_2 \,| > 0$ the Hessian $|\, H \,|$ is said to be *positive definite*, and the second-order conditions for the minimum are met. If $|\, H_1 \,| < 0$ and $|\, H_2 \,| > 0$ it is said to be *negative definite*, and the second-order conditions for the maximum are met.

§

Algorithm 1 *Procedure to test for the optimality of multivariable functions of two variables.*

$z = f(x, y)$
find z_x and z_y
if $z_x = 0$ and $z_y = 0$ **then**
  find z_{xx}, z_{xy} and z_{yy}
  find H_1 and H_2
  if $|\, H_1 \,| > 0$ and $|\, H_2 \,| > 0$ **then**
    $|\, H \,|$ **is positive definite**
  elseif $|\, H_1 \,| < 0$ and $|\, H_2 > 0 \,|$ **then**
    $|\, H \,|$ **is negative definite**
  endif
endif

Definition 5. Let $y = f(x_1, \ldots, x_n)$ be function of n variables. Then the n^{th}-order Hessian for this function is

$$|\, H \,| = \begin{vmatrix} y_{11} & \cdots & y_{1n} \\ \vdots & \ddots & \vdots \\ y_{n1} & \cdots & y_{nn} \end{vmatrix}$$

Then the *first principal minor* $|\, H_1 \,|$ is simply x_{11}, and the i^{th} principal minor is

$$|\, H_i \,| = \begin{vmatrix} y_{11} & \cdots & y_{1i} \\ \vdots & \ddots & \vdots \\ y_{i1} & \cdots & y_{ii} \end{vmatrix}$$

§

Theorem 4. Let $y = f(x_1, \ldots, x_n)$ be function of n variables. Let the Hessian of y be represented by $|\, H \,|$. Then if all the principal minors of $|\, H \,|$ are positive, then $|\, H \,|$ is positive definite and the second-order conditions for a relative minimum are met. If the sign of the principal minors alternates between negagive and positive, then $|\, H \,|$ is negative definite and the second-order conditions for a relative maximum are met.

§

Example 3. For $y = f(x_1, x_2, x_3)$ the third-order Hessian is

$$|\, H \,| = \begin{vmatrix} y_{11} & y_{12} & y_{13} \\ y_{21} & y_{22} & y_{23} \\ y_{31} & y_{32} & y_{33} \end{vmatrix}$$

where

$$y_{11} = \frac{\partial^2 y}{\partial x_1^2}, \quad y_{12} = \frac{\partial^2 y}{\partial x_2 \partial x_1}, \quad \text{and so on}$$

The first-, second- and third-order Hessian's are respectively

$$|\, H_1 \,| = y_{11}, \quad |\, H_2 \,| = \begin{vmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{vmatrix} \quad \text{and} \quad |\, H_3 \,| = \begin{vmatrix} y_{11} & y_{12} & y_{13} \\ y_{21} & y_{22} & y_{23} \\ y_{31} & y_{32} & y_{33} \end{vmatrix}$$

If $|H_1| > 0$, $|H_2| > 0$ and $|H_3| > 0$, then H is positive definite and the second-order condition for minimum is fulfilled. If $|H_1| < 0$, $|H_2| > 0$ and $|H_3| < 0$, then $|H|$ is negative definite and the second-order condition for maximum is satisfied.

Definition 6. A *discriminant* is a determinant of a quadratic form. Let the quadratic form be $z = ax^2 + bxy + cy^2$, which is in matrix form

$$z = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

Then the discriminant is

$$|D| = \begin{vmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{vmatrix}$$

The *first principal minor* of the discriminant is $|D_1| = a$, and the *second principal minor*

$$|D_2| = \begin{vmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{vmatrix} = ac - \frac{b^2}{4}$$

§

Theorem 5. Let a quadratic form be

$$z = ax^2 + bxy + cy^2$$

and let the discriminant of z be $|D|$. If $|D_1| > 0$ and $|D_2| > 0$, then $|D|$ is positive definite and $z > 0$ for all $x, y \neq 0$. If $|D_1| < 0$ and $|D_2| > 0$, then $|D|$ is negative definite and $z < 0$ for all $x, y \neq 0$.

§

Theorem 6. Let $f(x, y)$ be a function subject to a constraint $g(x, y) = k$, where k is a constant. Then the optimisation of f can be done by first transforming f together with g into a new function

$$F(x, y, \lambda) = f(x, y) + \lambda (k - g(x, y))$$

and then solve the following equations,

$$\begin{aligned} F_x(x, y, \lambda) &= 0 \\ F_y(x, y, \lambda) &= 0 \\ F_z(x, y, \lambda) &= 0 \end{aligned}$$

to obtain the critical values x_0, y_0 and λ_0 at which F and hence f are optimised.

§

Problem 4. Prove Theorem 5, the definiteness of a function by the discriminant. Prove the theorem for constrained optimisation with Lagrange multipliers, Theorem 6 above.

§

Definition 7. In the constrained optimisation with Lagrange multipliers in Theorem 6 above, f is called an *objective* or *origin function* and F the Lagrangian function.

§

Definition 8. Let $f(x_1, \ldots, x_n)$ be a function of n variables subject to constraints $g(x_1, \ldots, x_n)$. Let $F(x_1, \ldots, x_n, \lambda) = f(x, \ldots, x_n) + \lambda (k - g(x_1, \ldots, x_n))$. Then the *bordered Hessian* $|\bar{H}|$ is defined as either

$$|\bar{H}| = \begin{vmatrix} F_{11} & F_{12} & \cdots & F_{1n} & g_1 \\ F_{21} & & & & g_2 \\ \vdots & & \ddots & & \vdots \\ F_{n1} & & & F_{nn} & g_n \\ g_1 & g_2 & \cdots & g_n & 0 \end{vmatrix}$$

or

$$|\bar{H}| = \begin{vmatrix} 0 & g_1 & \cdots & g_n \\ g_1 & F_{11} & & F_{1n} \\ \vdots & & \ddots & \vdots \\ g_n & F_{n1} & \cdots & F_{nn} \end{vmatrix}$$

This is simply the Hessian

$$\begin{vmatrix} F_{11} & \cdots & F_{1n} \\ \vdots & \ddots & \vdots \\ F_{n1} & \cdots & F_{nn} \end{vmatrix}$$

bordered by the first derivatives of the constraint with zero on the principal diagonal. The order of a bordered principal minor being determined by the order of the principal minor being bordered, $|\bar{H}| = |\bar{H}_n|$ since in this case an $n \times n$ principal minor is being bordered.

§

Theorem 7. Let $f(x_1, \ldots, x_n)$ be a function of n variables subject to constraints $g(x_1, \ldots, x_n)$. Let $|\bar{H}|$ be the bordered Hessian defined in Definition 8. Then if $|\bar{H}_2|, \ldots, |\bar{H}_n| < 0$, then the bordered Hessian $|\bar{H}|$ is positive definite, and therefore is a sufficient condition for a minimum. If $|\bar{H}_2| > 0$ $|\bar{H}_3| < 0$, $|\bar{H}_4| > 0$, and so alternatingly on, then $|\bar{H}|$ is negative definite, which is a sufficient condition for a maximum.

§

Problem 5. Prove Theorem 7.

§

Example 4. Let $f(x, y)$ be a function to be optimised subject to a constraint $g(x, y) = k$, where k is a constant. Then the Lagrangian function becomes

$$F(x, y, \lambda) = f(x, y) + \lambda (k - g(x, y))$$

The first-order conditions for optimisation are $F_x = F_y = F_\lambda = 0$. The second-order conditions for optimisation can be expressed together as a bordered Hessian

$$|\bar{H}| = \begin{vmatrix} F_{xx} & F_{xy} & g_x \\ F_{yx} & F_{yy} & g_y \\ g_x & g_y & 0 \end{vmatrix}$$

or

$$|\bar{H}| = \begin{vmatrix} 0 & g_x & g_y \\ g_x & F_{xx} & F_{xy} \\ g_y & F_{yx} & F_{yy} \end{vmatrix}$$

Note 1. Theorem 6 gives the first-order conditions for optimising a function subject to some constraints. Theorem 7 gives the second-order conditions for optimising a function subject to some constraints.

§

Definition 9. A *Marshallian demand function* gives an expression of the amount of a good that a consumer will buy as a function of commodity prices and income available. It is derived by maximising the utility subjected to a budgetary constraint.

§

Example 5. Let a utility be $u = q_1 q_2$ which is subject to a constraint $p_1 q_1 + p_2 q_2 = b$, where b is the amount of income available, that is to say, our budget. Then the Lagrangian function is

$$U = q_1 q_2 + \lambda (b - p_1 q_1 - p_2 q_2)$$

The first partial derivatives are then

$$u_1 = q_2 - \lambda p_1 = 0 \tag{1}$$
$$u_2 = q_1 - \lambda p_2 = 0 \tag{2}$$
$$u_\lambda = b - p_1 q_1 - p_2 q_2 = 0 \tag{3}$$

where u_1, u_2 are respectively u_{q_1} and u_{q_2}. Simultaneously solving Equation's 1, 2 and 3 leads us to

$$\frac{q_2}{p_1} = \lambda = \frac{q_1}{p_2}$$

Hence $q_2 = q_1 p_1/p_2$ and $q_1 = q_2 p_2/p_1$ and from Equation 3 we have,

$$b = p_1 q_1 + p_2 \frac{p_1 q_1}{p_2} = p_2 q_2 + p_1 \frac{p_2 q_2}{p_1}$$

which yield us, for q_1 and q_2, the Marshallian demand functions which maximise satisfaction of the consumer subject to income and prices.

Next, we test the second-order conditions by firstly finding $u_{11} = 0$, $u_{22} = 0$, $u_{12} = u_{21} = 1$, $g_1 = p_1$ and $g_2 = p_2$, which give us

$$|\bar{H}| = \begin{vmatrix} 0 & 1 & p_1 \\ 1 & 0 & p_2 \\ p_1 & p_2 & 0 \end{vmatrix}$$

which gives $|\bar{H}_2| = 2p_1p_2 > 0$ Hence $|\bar{H}|$ is negative definite and thus u is maximised.

Definition 10. The production process of producing one good usually requires the input of many other *intermediate goods*. Let x_i be the total demand for product i, and let b be the final demand for the product from the ultimate users. Then,

$$x_i = a_{i1}x_1 + \ldots + a_{in}x_n + b_i$$

for $i = 1, \ldots, n$, where a_{ij} is a *technical coefficient* which represents the value of input i required to produce one monetary unit's worth of product j. If we consider the total demand for every one of the products, then

$$\mathbf{x} = A\mathbf{x} + \mathbf{b}$$

where

$$\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \quad A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$$

It follows from this that $\mathbf{x} = (I - A)^{-1}\mathbf{b}$. The matrix A is known as the *matrix of technical coefficients*. It is also known as the *input-output table*, the rows being the inputs and the columns the outputs.. The matrix $I - A$ is known as the *Leontief matrix*.

§

Example 6. In a complete input-output table, labour and capital would also be included as inputs. These give the value added by the firm. They are normally put as an extra row at the bottom of the matrix of technical coefficients A. The vertical summation of each column of the table is then equal to 1.

Definition 11. Let A be a square matrix. Then a scalar λ such that the equation

$$A\mathbf{v} = \lambda\mathbf{v} \tag{4}$$

holds for some vector $\mathbf{v} \neq \mathbf{0}$ is called an *eigenvalue* † of A, and the vector $\mathbf{v}$ is called an *eigenvector* of A corresponding to the eigenvalue λ. The eigenvalue λ is also known as the *characteristic root*, or the *latent root*, while the eigenvector is also known as the *characteristic vector*, or the *latent vector*.

§

Note 2. From Equation 4 it follows directly that

$$(A - \lambda I)\mathbf{v} = 0 \tag{5}$$

† The word *eigenvalue* is a half-translation of the German word *Eigenwert*. The latter means 'appropriate value' since *Wert* means 'value' and *eigen* means 'proper or appropriate'.

Then $A - \lambda I$ is called the *characteristic matrix* of A. Since $\mathbf{v}$ assumes a unique value and by assumption $\mathbf{v} \neq 0$, it follows that $A - \lambda I$ must be singular, which means that its rows must be a multiple of one another. Now $A - \lambda I$ is zero if and only if the *characteristic determinant* $|A - \lambda I|$ of A is zero. In other words

$$|A - \lambda I| = 0 \tag{6}$$

which is called the *characteristic equation* of A. With Equation 6 there will be an infinite solution for $\mathbf{v}$ in Equation 5. In particular, if $\mathbf{v}$ is a solution, that is if it is an eigenvector, so is $k\mathbf{v}$ for any $k \neq 0$. We force a unique solution by using the *normalisation*

$$\sum v_i^2 = 1$$

Then the sign-definiteness of A can be determined from the characteristic roots λ's. Thus if all λ's are positive, then A is positive definite; and if negative, negative definite. Let at least one λ be zero, which is neither positive nor negative, if all the remaining λ's are nonnegative, then A is positive semidefinite; and if they are nonpositive, negative semidefinite. Lastly, if some of the λ's are positive while others are negative, then A is indefinite.

§

Problem 6. Prove all the necessary details need to be proved in Note 2.

§

Note 3. We have seen in Note 2 how, having found λ_i, where $i = 1, \ldots, n$, we find through normalisation the corresponding, unique $\mathbf{v}_i$. On the other hand if we have found first the $\mathbf{v}_i$'s, their corresponding λ_i's may be found by first forming a *transformation matrix*

$$T = [\mathbf{v}_1 \ \ldots \ \mathbf{v}_n]$$

and then the corresponding eigenvalues or the characteristic roots are obtained from

$$T^T A T = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & \\ \vdots & & \ddots & \vdots \\ 0 & & \cdots & \lambda_n \end{bmatrix}$$

§

Problem 7. Prove all the necessary details in Note 3.

§

Definition 12. The vector equation, Equation 4, has as its solutions the zero vector $\mathbf{v} = 0$ together with all the corresponding eigenvalue-eigenvector pairs. The set of all the eigenvalues of A is called the *spectrum* of A. The *spectral radius* of A is then the largest of all the absolute values of the eigenvalues of A, that is to say,

$$\max_i |\lambda_i|$$

The set of all eigenvectors $\mathbf{v}_{ij}$, together with $\mathbf{0}$, forms a vector space called the *eigenspace* of A corresponding to λ_i.

§

Bibliography

Edward T Dowling. *Introduction to mathematical economics.* Schaum's outline series,2nd ed. 1992(1980)

Erwin Kreyszig. *Advanced engineering mathematics.* 7th ed, 1993